Sun StorEdge™ NAS Gateway System
Configuration and Usage Guide

White Paper
August 2005

Table of Contents

Introduction

The Sun StorEdge™ 5310 NAS Gateway System (SE5310-G) provides NAS services to both NFS and CIFS users, drawing on block storage allocated from a pooled SAN storage resource. A clustering option is available, allowing a pair of StorEdge 5310 Clustered Gateway Systems (SE5310-CG) to operate as an active-active failover protected set, providing both increased availability and increased performance.

Network Attached Storage Concepts

Unlike a SAN, where pooled storage resources are allocated to specific servers on an exclusive basis, the goals of a NAS service are to provide file-based storage services for information which may be either private to a particular user, or shared broadly and heterogeneously across a wider community of authorized participants. To accomplish this, Network Attached Storage relies on a client-independent implementation of a file service, allowing the same variable-sized data objects to be accessed regardless of the client machine's type or its operating system version.

These file services are accessed using standards-based protocols such as Network File System (NFS), Common Internet File System/Server Message Block (CIFS/SMB), and File Transfer Protocol (FTP). Each of these protocols embodies the concept of a “user”, an authenticated individual having certain access rights and privileges to information stored by the file service. In general, the NFS protocol is most commonly seen in Unix and Linux user communities, while the CIFS/SMB protocol is more frequently seen in Windows environments. The Sun StorEdge 5310 Gateway System supports both protocols simultaneously, as well as the legacy FTP service.

The file system used by the Sun StorEdge 5310 NAS Gateway System supports file attributes that can be mapped into both “Unix” and “NTFS” representations, as used by NFS and CIFS respectively. Similarly, the Sun StorEdge 5310 NAS Gateway System supports two families of user authentication services; NIS, NIS+, and LDAP for the <UID,GID> user representation used by NFS, and Windows Domain and Active Directory for the username-based credentials used by CIFS. The Sun StorEdge 5310 Gateway System may be configured to map between these two user models, allowing files to be shared between Unix and Windows user communities.

The file system also supports both hard and soft quotas, which may be assigned to a file system volume on a per user or per user group basis. Quotas may also be assigned to directory trees within a given file system.

Network Interface

Attachment to a local area network (LAN) infrastructure provides the NAS server with a scalable, sharable interface to the user community. The Sun StorEdge 5310 NAS Gateway provides two 10/100/1000 Base-T Network connections as part of its server motherboard, and can be fitted with two additional Optical Gigabit Ethernet ports as an option.

The primary purpose of these interfaces is to provide connectivity for NAS services between a client community and the Sun StorEdge 5310 Gateway System and to support administrative access for monitoring and configuration. At a minimum, one interface per server must be assigned an IP address and be attached to a LAN infrastructure providing access to the client community using the desired NAS protocols.

Enabling multiple LAN ports improves connectivity and increases performance. This may be done in two ways; by assigning different IP addresses to additional ports, or by creating a multiport link aggregation group sharing a single IP address (also described as a “port aggregation” or “Etherchannel” bond.) Each alternative has its own unique constraints.

The Sun StorEdge 5310 NAS Gateway System network stack relies on a single gateway (IP router) address for all ports and each interface that has a distinct IP address must be on a separate IP subnet. Otherwise, the external network infrastructure would be unable to forward return messages properly to their originating server port.

If link aggregation is enabled, all ports in the group share a single IP address and thus are part of the same subnet. The link aggregation group's termination on a LAN switch must be configured to enable trunking support. For current LAN switches, this in general means that all ports in a link aggregation group must terminate on the same physical switch (or, for some LAN switches, on certain designated ports of the switch.) Because few Ethernet switches support mixtures of Copper and Optical physical interfaces, this translates in practice to a de facto requirement that all ports in a link aggregation group be of the same type, either copper or fiber.

Unless direct connectivity to multiple subnets is required because of network design, the preferred practice is to enable trunking across two ports of identical type and speed, attached to ports on the same Ethernet Switch via equivalent paths (i.e. similar length, cable material, number of junction points, etc.)

It should be noted that one mechanism commonly used in trunking implementations to distribute traffic among the individual links in an aggregation is a static function of source and destination MAC address. Where a few application sessions are expected to generate the bulk of the NAS activity, they should be spread across multiple client systems to prevent overloading of individual links in an aggregation group.

Impact of Clustering on Network Interfaces

A Sun StorEdge 5310 NAS Clustered Gateway System consists of two identical servers. In normal operation, each server operates independently, offering their own set of NAS shares to the client community and thus each responds only to client requests for their own set of services. Each server also monitors its peer's health by exchanging status messages across a private point-to-point connection. These messages also communicate configuration changes and significant operational events (such as system shutdown requests) to the peer system.

For clustered servers using the onboard copper Ethernet ports as their primary NAS attachments, the health check connection is provided by a single port 10/100 Base-T NIC, which is a standard addition to this configuration. Clustered servers built with the additional Optical Gigabit NIC are assumed to rely on those ports for primary NAS connection, and instead utilize one of their motherboard copper ports for the health monitoring function, thus not requiring addition of the 10/100Base-T NIC. In either case, the *recommended configuration* reserves the designated ports for heath monitoring, using a direct point-to-point (crossover) Ethernet cable between the two systems' health monitor ports rather than LAN connection. Use of the designated health monitoring port for other purposes is not supported.

On a server failure, typically detected by failure to respond to a health check, the surviving server enters a fault recovery mode. In this mode, it adds the IP address(es) of its failed peer to its own LAN interfaces, mounts its peer's LUN volumes, and begins to offer its peer's NAS shares as well as its own. Subsequent client connections to these services will be directed to the surviving server now issuing the share offers, and existing connections will transition as incomplete client requests are retried.

This failover operation between servers imposes additional design requirements on network connectivity for a Sun StorEdge 5310 NAS Clustered Gateway System. The surviving server must be able to successfully operate using the IP addresses of its failed peer, therefore, both servers in a cluster are required to have interfaces on the same set of subnets. The transfer of network addresses between servers during a failover also requires that the LAN switches not be set in a “lock down security” mode, which would prohibit them from forwarding traffic based on the new relationship of addresses and physical ports seen after a failover.

If Etherchannel port aggregation or High Availability port bonding is used, it is *strongly recommended* that both servers in the cluster have identical LAN port configurations. This will both simplify configuration, and help minimize the potential performance impact of a failover by providing comparable network bandwidth to both servers.

Storage Attachment

Unlike the Sun StorEdge 5310 NAS Appliances that rely on privately owned storage, Sun StorEdge 5310 NAS Gateway Systems have much sharper delineation between their File Server and Block Storage components. Each Sun StorEdge 5310 NAS Gateway System performs NAS protocol processing and maintains file systems, while external storage in the SAN handles all drive management, storage redundancy, and LUN mapping. Each component (SAN and NAS) handles its own configuration, event recording, and failure recovery, reporting only general status rather than specific detail to the other system components.

The Sun StorEdge 5310 NAS Gateway System may be attached directly to a multiport array, but in most cases a switched SAN fabric will be used for connectivity. The type of fabric switch used and their supported topology must be supported by both the Sun StorEdge 5310 NAS Gateway System and the storage array being used.

Note: Once a Sun StorEdge 5310 NAS Gateway System has been configured (with either direct attachment or SAN fabric attachment) the attachment type may not be changed.

Storage Considerations for Clustered Systems

Prior to a failover event, the two servers in a Sun StorEdge 5310 NAS Clustered Gateway System are independent, offering separate sets of storage shares to NAS clients, and drawing upon separate LUN volumes for file system storage. These volumes may be on separate arrays or on a common storage system, the only requirement being that each server have direct access to its storage subsystem.

During failover, the surviving server in the cluster must successfully mount the storage volumes used by its peer, as well as its own volumes. To support a cluster failover, each server in a Sun StorEdge 5310 NAS Clustered Gateway System must not only have redundant multi-path access to its own volumes, but to all volumes used by itself and its peer.

Storage Usage

Each Sun StorEdge 5310 NAS Gateway System must have its own storage resources, provided as dedicated LUNs by the storage arrays. A minimum of one LUN per Sun StorEdge 5310 NAS Gateway System must be provided, with up to 64 LUNs per server supported.

Note: For a Sun StorEdge 5310 NAS Clustered Gateway System, the number of LUNs per server must be constrained such that, when a failover occurs, the surviving server will have no more than 64 LUNs.

The Sun StorEdge 5310 NAS Gateway System allows dynamic growth of its file systems by drawing upon multiple block volumes through its storage segment concept. The Sun StorEdge 5310 NAS Gateway System divides its LUNs into segments of up to 256 GB each. One such segment is used as the initial storage (primary segment) for creation of a file system. Subsequently, up to 63 additional segments may be added to that file system, allowing it to dynamically grow to as much as 16 TB capacity. Creating file systems comprised of multiple segments from different LUNs provides similar advantages to broader striping of the underlying storage volumes, in that file system operations are spread

across a greater number of disk drives. However, unlike a single very wide RAID parity group, the multiple RAID groups combined using the segment concept retain their relatively fast fault reconstruction times.

Each LUN allocated to a Sun StorEdge 5310 NAS Gateway System must be accessible via two distinct paths, to provide continued access in the event of a port, link, or path failure. This is true both for direct attachment, where the redundant paths would consist of separate fibre channel connections from the fibre channel HBA on the Sun StorEdge 5310 NAS Gateway System to two storage array controller ports, and for SAN fabric attachment, where the concept of redundant paths extend to separate fabric connections and, if required by site SAN guidelines, independent fabric switches.

Both the storage array and the SAN infrastructure must be configured to allow each LUN assigned to Sun StorEdge 5310 NAS Gateway System use to be accessible via either path. Fabric switches must be zoned such that access is limited to only the Sun StorEdge 5310 NAS Gateway System and the storage array ports assigned to it. With a Sun StorEdge 5310 NAS Clustered Gateway System, the zone must extend to allow any port on either server to access any LUN assigned to either server's use.

Supported Configurations

Note: The configuration illustrations following depict Sun StorEdge 99xx Systems. Similar configurations utilizing the Sun StorEdge 6920 are also supported.

Direct Attach for Single Systems

The minimum supported configuration for a Sun StorEdge 5310 NAS Gateway system is illustrated in Figure 1 below and depicts two fibre channel connections between the server and the storage array. From the server a single port from each fibre channel HBA is utilized and on the storage array a pair of high-availability ports is utilized. In this configuration all LUNs are available to all ports and if any given port fails there is an alternate path available for the data access.

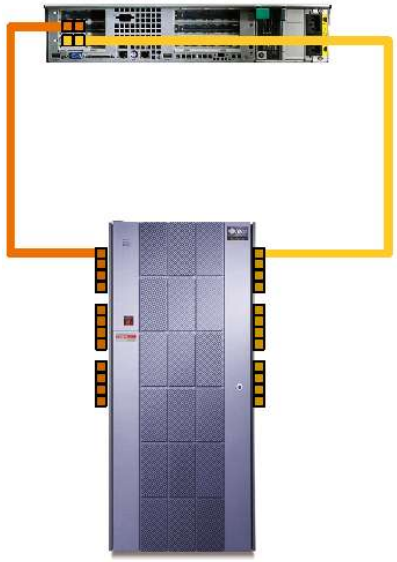
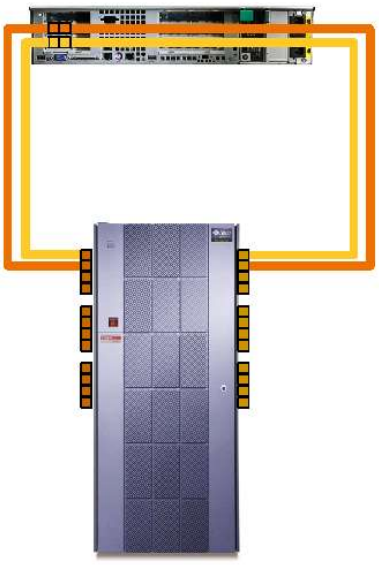


Figure 1. Single system with direct attachment utilizing a single pair of storage array ports

To provide higher bandwidth, an additional pair of ports may be utilized on the storage array as illustrated in Figure 2 below. This configuration utilizes all the available ports from the two fibre channel HBAs on the system and two separate pairs of high availability ports on the storage array. In this configuration, each pair of high availability ports on the storage array have their own LUNs and the LUNs available from both pairs are visible to the system. Furthermore, if any given path fails, their remains a surviving path for data access.

Figure 2. Single system with direct attachment utilizing two pairs of storage array ports



Direct Attach for Clustered Systems

When configuring a Sun StorEdge 5310 NAS Clustered Gateway System with direct connection, the minimum configuration requires two pairs of high availability ports on the storage array as illustrated in figure 3 below. In this configuration all LUNs must be shared between all the ports in the storage array to allow for a system failover in addition to a potential path failure.

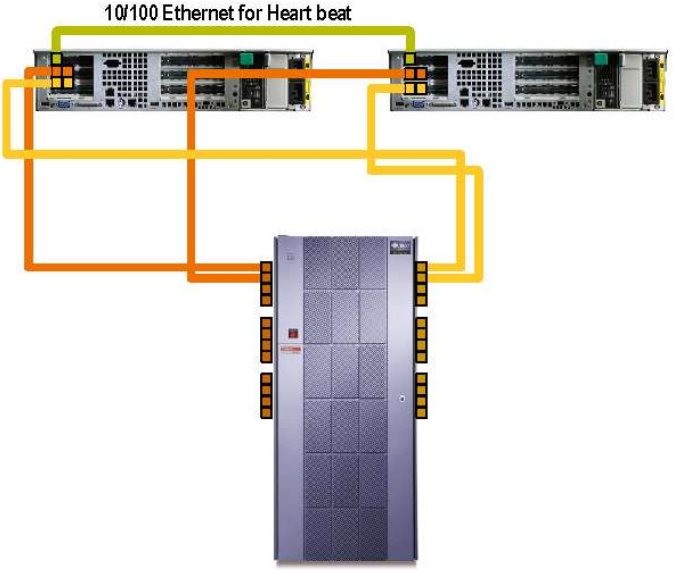


Figure 3. Clustered system with direct attachment utilizing two pairs of storage array ports

For additional bandwidth, the configuration illustrated in Figure 4 below may be implemented which utilizes all four ports on each server and four pairs of high availability ports on the storage array. In this configuration as well, all LUNs must be shared between all the ports in the storage array to allow for a system failover in addition to a potential path failure.

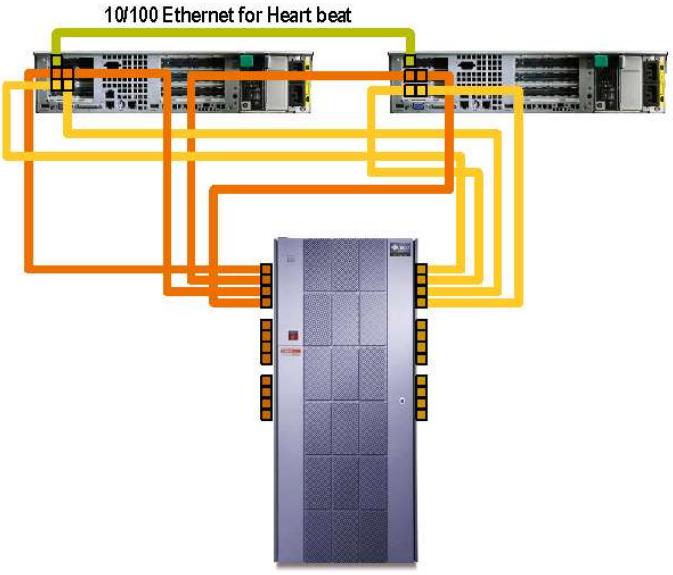


Figure 4. Clustered system with direct attachment utilizing four pairs of storage array ports

Fabric Attach

Note: When configuring Sun StorEdge NAS5310 Gateway systems with a SANSwitch, the maximum cascade level is 2.

Sun StorEdge NAS5310 Gateway systems can also be configured utilizing SANSwitches. In all configurations it is highly recommended that two SANswitches be used to ensure there is no single point of failure in the data path.

Configurations utilizing fabric attach are all similar to direct attach in as much as the LUN/Port associations of the storage array remain the same.

For all illustrations using SANswitches the ports on the SANswitches are color coded with each color indicating a unique Zone (e.g. Zone 1 and Zone 2) as illustrated below in Figure 5.

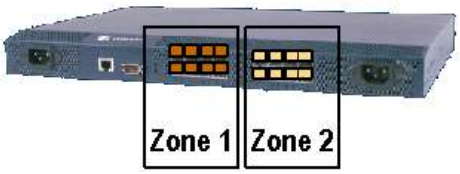
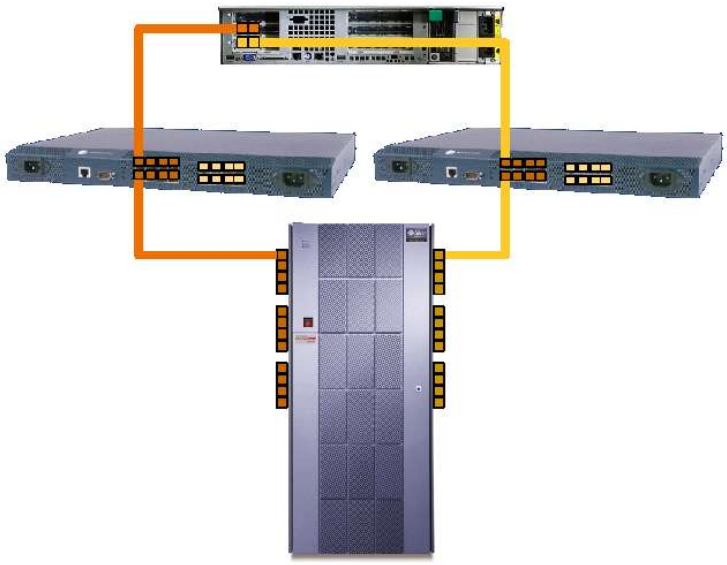


Figure 5. Illustration of SAN Switch Zoning

Fabric Attach for Single Systems

The minimum recommended configuration for a Sun StorEdge 5310 Gateway System attached via fabric is illustrated in Figure 6 below.

Figure 6. Single system with fabric attachment utilizing a single pair of storage array ports



When utilizing fabric attach, multiple (non-clustered) systems can also be configured to utilize the same ports on the storage array as illustrated in Figure 7 below. In this configuration the LUNs are shared between the ports on the storage array but allocated uniquely to each system.

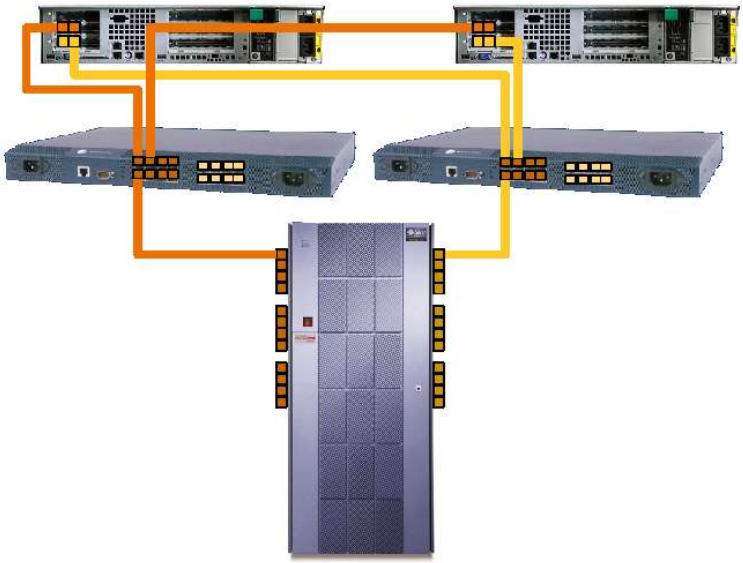


Figure 7. Multiple single systems with fabric attachment utilizing a single pair of storage array ports

slight advantage in terms of path contention, separate zoning not required,

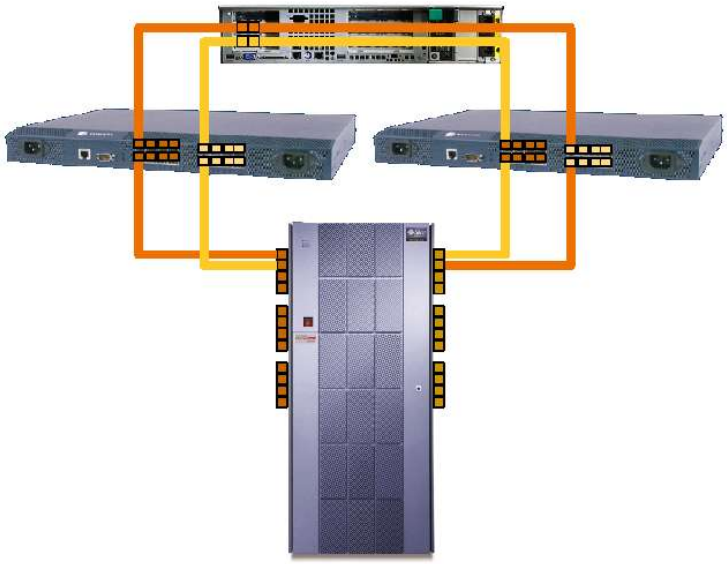


Figure 8. Single system with fabric attach utilizing four system ports and a two pairs of storage array ports

Fabric Attach for Clustered Systems

The minimum configuration for Sun StorEdge 5310 NAS Clustered Gateway Systems requires that a single port from each server be connected via dual switches utilizing a single pair of high availability ports on the storage subsystem as illustrated in Figure 9 below. This allows for the failure of any given port, switch or server with the Clustered System while still providing available paths to the data.

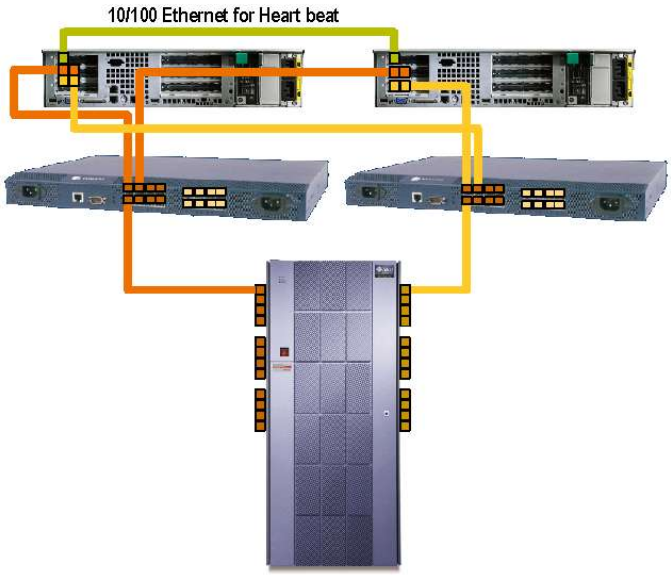


Figure 9. Clustered system with fabric attach utilizing two system ports and a single pair of storage array ports

To provide higher bandwidth from the SAN storage array, additional pairs of high availability ports may also be used to balance against the connectivity from the Sun StorEdge 5310 NAS Clustered Gateway System servers as illustrated in Figure 10 below.

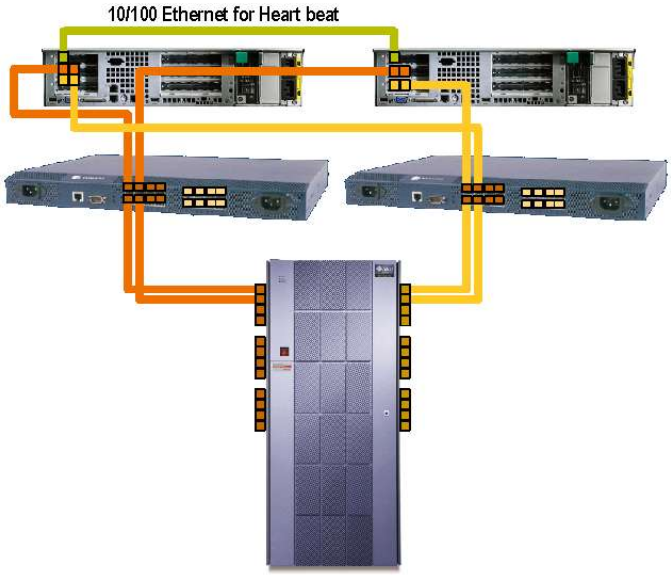


Figure 10. Clustered system with fabric attach utilizing four system ports and two pairs of storage array ports

The maximum supported fabric connected configuration is illustrated in Figure 11 below which depicts the usage of all

four available ports from the Sun StorEdge 5310 NAS Clustered Gateway System servers and up to four pairs of high availability ports on the Sun StorEdge 9970, 9980 or 6920 Systems.

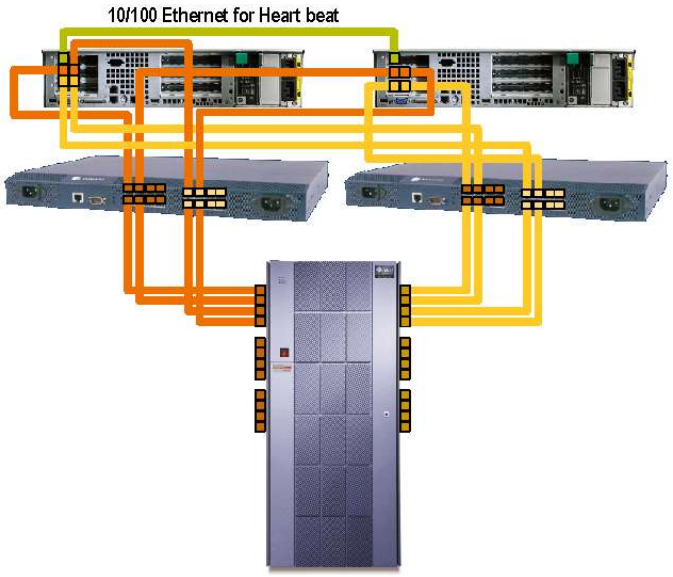


Figure 11. Clustered system with fabric attach utilizing four system ports and four pairs of storage array ports

Installing the Sun StorEdge 5310 NAS Gateway System

For detailed installation instructions for the Sun StorEdge 5310 NAS Gateway System please refer to both the Hardware User Guide and the Software Installation and Administration Guide.

Configuring the Sun StorEdge 99xx System for the Sun StorEdge 5310 NAS Gateway System

The following are required configuration parameters for Sun StorEdge 9970/9980 Systems for use with the Sun StorEdge 5310 NAS Gateway System:

Host Mode – 00 (standard)

Port connection – FC-AL

Fabric - ON

Both direct attachment and switched fabric attachment between Sun StorEdge 5310 NAS Gateway System and the Sun StorEdge 99xx System is supported. If fabric attachment is used, zoning must be used to prevent hosts other than Sun StorEdge 5310 Gateway System from obtaining access. The zone used with Sun StorEdge 5310 NAS Clustered Systems (as shown in the previous figures) must extend to include both servers in the cluster, and all Sun StorEdge 9900 System ports assigned to that NAS cluster's use.

Device Types and Configuration Procedures for the Sun StorEdge 9900 System

The Sun StorEdge 9900 System allows the different types of logical devices (also called volumes) to be installed and configured for operation with the Sun StorEdge 5310 NAS Gateway System these are outlined in Table 1 below.

Note: See the configuration guide for your SE9900 subsystem for device specifications and recommended volume usage for these logical devices.

Table 1. Valid Sun StorEdge 9900 System Device Types

Device Type	Description
Virtual LVI/LUN Devices (OPEN-V)	Virtual LVI/LUN (VLL) allows you to configure variable-size volumes, which are usually smaller than normal (fixed-size) volumes. Virtual LVI/LUN improves data access performance by reducing logical device contention and host I/O queue times, particularly when several frequently accessed files are located on the same volume. Virtual LVI/LUN enables better utilization of the physical storage capacity of the 9900V, and reduces the amount of administrative effort required to balance I/O workloads. For further information, please refer to the *Hitachi Lightning 9900TM V Series LUN expansion and Virtual LVI/LUN User's Guide.*
VIR LUSE Devices (OPEN-*x**n VIR).	The VIR LUSE devices combine VIR devices (instead of standard OPEN-*x* LUs) into LUSE devices. The VIR feature is used to create custom-size devices, and then the LUSE feature is used to combine (concatenate) these VIR devices. The user can combine from 2 to 36 VIR devices into one VIR LUSE device. For example, an OPEN-3 LUSE volume that was created from ten OPEN-3 VIR volumes would be designated as an OPEN-3*10 VIR device
OPEN-*x* Devices.	The OPEN-*x* logical units (LUs) (e.g., OPEN-3, OPEN-9) are disk devices of predefined sizes. The 9900 subsystem currently supports OPEN-3, OPEN-8, OPEN-9, OPEN-K, and OPEN-E, OPEN-L and OPEN-M devices. Please contact your Hitachi Data Systems account team for the latest information on supported LU types.
LUSE Devices (OPEN-*x**n).	The LUSE devices are combined LUs which can be from 2 to 36 times larger than standard OPEN-*x* LUs. The Logical Unit Size Expansion (LUSE) feature of the 9900 subsystem enables you to configure these custom-size devices. LUSE devices are designated as OPEN-*x**n, where *x* is the LU type (e.g., OPEN-9*n) and $2 < n < 36$. For example, a LUSE device created from ten OPEN-3 LUs would be designated as an OPEN-3*10 disk device. This capability enables the server host to combine logical devices and access the data stored on the 9900 subsystem using fewer LU numbers (LUNs). For further information on the LUSE feature, please refer to the *Hitachi Lightning 9900TM LUN Manager User's Guide* (MK-91RD049).

The Sun StorEdge 5310 NAS Gateway System will format the logical device as a file system and Sun recommends use of OPEN-V Virtual LVI/LUN volumes for this purpose. OPEN-V volumes of any size up to 737,256 MB may be created.

Note: Defining a VLL volume larger than 60 gigabytes automatically creates the VLL volume in LUSE configuration mode.

If LUSE devices are used with the Sun StorEdge 5310 NAS Gateway System, it is recommended that each LUSE be within a parity group so that striping across multiple LUSE will also stripe across multiple parity groups.

Configuring the Sun StorEdge 6920 System for the Sun StorEdge 5310 NAS Gateway System

The interconnection between the SE5310-G or -CG and a SE6920 storage system may be either direct attach, or fabric attach. A minimum of two paths must be provided, to allow for module and port failover. Up to four paths may be used, and are recommended if a large number of LUNs are configured for use by the NAS Gateway.

To simplify configuration of SE6920 storage for use with the SE5310 NAS Gateway, two storage profiles have been defined appropriate for NAS use. They are:

NFS_Stripe – recommended for general NAS usage, and read-intensive operations

NFS_Mirror – recommended for write-intensive NAS applications

Configuring the Sun StorEdge 5310 NAS Gateway System

For instructions on configuring the Sun StorEdge 5310 NAS Gateway system, please refer to the Sun StorEdge 5310 *Software User Guide* chapters 1-4 for initial network, user, and storage configuration.

Restrictions

Known restrictions at this time are the following:

- Only one Sun StorEdge 9900 System may be allocated per Sun StorEdge NAS Gateway System installation (single or clustered pair)
- Supported Switches and topologies
  - Brocade Silkworm 3850/3250 Firmware revision level 4.2.2
  - Maximum 2 tiers of fabric switching,
- WWN Switch Zoning required
- No Changes from Direct to Fabric Attach after installation
- No Extended features may be used with Sun StorEdge 9900 Systems
  - Replication
  - Snap
  - Etc.
- Two paths are required from each Sun StorEdge 5310 Gateway system to its LUNs (to all LUNs used by both systems in a cluster configuration)
- Sun StorEdge Compliance Archiving Software is not supported
- No LUN movement from one Sun StorEdge 5310 Gateway System to another Sun StorEdge 5310 Gateway System

Sun Microsystems, Inc. Network Circle, Santa Clara, CA 95054 USA Phone 1-650-960-1300 or 1-800-555-9SUN Web sun.com

© 2005 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, CA 95054 USA
All rights reserved.
This product or document is protected by copyright and distributed under licenses restricting its use, copying, distribution, and decompilation. No part of this product or document may be reproduced in any form by any means without prior written authorization of Sun and its licensors, if any. Third-party software, including font technology, is copyrighted and licensed from Sun suppliers.

Parts of the product may be derived from Berkeley BSD systems, licensed from the University of California.

Sun, Sun Microsystems, the Sun logo, [ADD APPLICABLE TRADEMARKS HERE] are trademarks, registered trademarks, or service marks of Sun Microsystems, Inc. in the U.S. and other countries.

UNIX is a registered trademark in the United States and other countries, exclusively licensed through X/Open Company, Ltd.
All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc.

The OPEN LOOK and Sun™ Graphical User Interface was developed by Sun Microsystems, Inc. for its users and licensees. Sun acknowledges the pioneering efforts of Xerox in researching and developing the concept of visual or graphical user interfaces for the computer industry. Sun holds a non-exclusive license from Xerox to the Xerox Graphical User Interface, which license also covers Sun's licensees who implement OPEN LOOK GUIs and otherwise comply with Sun's written license agreements.

RESTRICTED RIGHTS: Use, duplication, or disclosure by the U.S. Government is subject to restrictions of FAR 52.227-14(g)(2)(6/87) and FAR 52.227-19(6/87), or DFAR 252.227-7015(b)(6/95) and DFAR 227.7202-3(a). DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS HELD TO BE LEGALLY INVALID.